

**EUROPEAN PATENT APPLICATION**

Application number: **88301649.5**

Int. Cl.<sup>4</sup>: **H03D 3/00** , **H03D 1/00**

Date of filing: **24.02.88**

Priority: **27.02.87 GB 8704871**

Date of publication of application:  
**05.10.88 Bulletin 88/40**

Designated Contracting States:  
**DE FR GB IT NL**

Applicant: **BRITISH AEROSPACE PUBLIC  
LIMITED COMPANY**  
**11 Strand**  
**London WC2N 5JT(GB)**

Inventor: **Biggs, Roger Tredegar**  
**c/o Box 115, British Aerospace PLC**  
**Downshire Way**  
**Bracknell Berkshire, RG12 1QL(GB)**  
Inventor: **Standford, Richard Thomas Albert**  
**c/o Box 115, British Aerospace PLC**  
**Downshire Way**  
**Bracknell Berkshire, RG12 1QL(GB)**

Representative: **Rackham, Stephen Neil et al**  
**GILL JENNINGS & EVERY 53-64 Chancery**  
**Lane**  
**London WC2A 1HN(GB)**

**Digital signal demodulation.**

A method of demodulating a high frequency carrier signal having at least one parameter modulated by a low frequency modulating signal comprises translating the signal to a low intermediate frequency, digitising the translated intermediate frequency signal at a correspondingly low sampling rate and applying terms from the resulting digital time series to a non-linear processor to produce an output representative of the instantaneous value of the modulated parameter. By deriving an instantaneous value of the modulated parameter removes the need to use a high intermediate frequency for the carrier and enables the signal to be sampled at a low sampling rate.

**EP 0 285 252 A2**

### Digital Signal Demodulation

In conventional analogue modulation techniques a signal is used to modulate one or more of the parameters of a carrier wave i.e. phase, amplitude or frequency. The carrier wave is usually of a frequency significantly greater than that of the modulating signal. As a result the modulated parameter may be assumed to be constant over several cycles of the carrier wave. Conventional demodulation techniques, such as zero-crossing counting in the case of frequency modulation, make use of this to obtain accuracy by averaging over several carrier signal cycles.

In applying digital signal processing techniques to demodulation the methods used to date have in general been functionally equivalent to the standard analogue techniques. Examples of such techniques are described in GB-A-2144288; EP-A-0046213; and, US-A-4567442. Typically the carrier is translated to an intermediate frequency by means of a superheterodyne process but this intermediate frequency still has to be high enough for the assumption that the modulated parameter is constant to hold for most demodulation techniques. In the case of frequency modulation, for example, digital zero-crossing techniques have been used in which the frequency at which the numerical representation of the signal falls below a threshold value is assessed over several cycles. Such techniques require a sampling rate high enough to produce an accurate digital representation of the modulated intermediate frequency signal since lower sampling rates result in unacceptable levels of distortion caused by aliasing. Therefore even though digitisation is carried out after translation of the carrier to the intermediate frequency it is still necessary to use a high sampling rate at least as great as the Nyquist rate characteristic of the intermediate frequency. The need to sample at a high rate and to transmit and process the resulting large bandwidth digital signal, a substantial part of which is redundant since all the information is contained within the narrower bandwidth characteristic of the modulating signal, is in many applications a significant disadvantage.

According to the present invention a method of demodulating a signal comprising a high frequency carrier signal having at least one parameter modulated by a low frequency modulating signal comprises translating the signal to a low intermediate frequency, digitising the translated intermediate frequency signal at a correspondingly low sampling rate and applying terms from the resulting digital time series to a non-linear processor to produce an output representative of the instantaneous value of the modulated parameter.

By deriving an instantaneous value of the modulated parameter rather than deriving a value over several cycles of the carrier a method in accordance with the present invention removes the need to use a high intermediate frequency for the carrier. With conventional demodulation techniques the intermediate frequency  $f_i$  has had to satisfy the relationship  $f_i \gg f_m$  where  $f_m$  is the highest frequency of the modulating signal. Thus a typical FM receiver operating in the broadcast band of 88 to 108 MHz uses a high intermediate frequency of 10.7 MHz. If the modulating signal has a maximum frequency  $f_m$  of 15 KHz the ratio of  $f_i$  to  $f_m$  is greater than 500:1. Digitising the signal would therefore require the use of a sampling rate at least as great as the Nyquist rate characteristic of the signal at the high intermediate frequency of 10.7 MHz. The present invention by contrast enables the use of a low intermediate frequency such that the ratio of  $f_i$  to  $f_m$  is significantly smaller than with known techniques: typically  $f_i$  may be of the same order of magnitude as  $f_m$ . The ability to use a correspondingly low sampling rate offers significant advantages in terms of cost and performance of components such as A/D converters. A low ratio between the intermediate frequency and the modulating frequency also confers important functional advantages in the many applications where it is necessary to minimise the redundant bandwidth of the digitised signal as, for example, when multiplexing a number of such signals for transmission over a channel of restricted bandwidth.

The non-linear processor may perform a Fast Fourier Transform to produce an output representative of the instantaneous value of the modulated parameter but preferably uses sinusoidal curve-fitting.

Preferably at least three successive terms of the digital time series are applied to the non-linear processor to produce the output representative of the instantaneous value of the modulated parameter. Preferably a method in accordance with the present invention further comprises taking one further term from the digital time series and applying the one further term and all but the first of the at least three successive terms to the non-linear processor to produce a further output representative of the next instantaneous value of the modulated parameter, and so on.

By substituting the at least three consecutive values of the digital time series in the equation of a sinusoidally varying signal it is possible to solve for any of the three parameters, phase, frequency or amplitude. This method of demodulation is therefore equally applicable to frequency modulation, phase modulation and amplitude modulation. This contrasts with the known methods of demodulation which are dedicated to the detection of a single form of modulation. Equally the initial translation to a low intermediate

frequency and sampling processes are applicable to phase, amplitude or frequency modulated signals. Moreover it is not necessary to remove any modulation of the other parameters before a given parameter is demodulated. This again contrasts with conventional demodulation methods: known methods of frequency demodulation, for example, require that the signal is limited to remove any significant amplitude modulation before it is applied to a frequency discriminator. These advantages make demodulation in accordance with the present invention particularly advantageous when applied to the demodulation of intercepted communications since it is not possible to predict in advance the nature of the modulation used in such signals.

A method of demodulation in accordance with the present invention is now described in detail with reference to the accompanying drawings, in which:-

Figure 1 is a flow diagram of a method in accordance with the present invention;

Figure 2 is a diagram illustrating the sampling processes used in a method in accordance with the present invention;

Figure 3 is a flow diagram of a program used to evaluate the performance of a method in accordance with the present invention; and

Figures 4 to 9 are graphs showing the results of an evaluation of the performance of a method in accordance with the present invention;

A radio frequency analogue signal comprising a high frequency carrier wave frequency modulated by a low frequency signal, is applied to the input of a demodulator 1. A superheterodyne circuit 2 translates the carrier wave to a low intermediate frequency  $f_i$  such that  $1.2 \leq f_i/f_m$  where  $f_m$  is the maximum frequency of the modulating signal: in practice a typical value is  $f_i/f_m \approx 10$ . The translated signal is then applied to an A/D converter 3 which, (as shown in Figure 2) samples the signal at a low sampling rate greater than the Nyquist rate characteristic of the low intermediate frequency  $f_i$  and outputs a digital time series in which each term represents an instantaneous amplitude of the translated signal. Groups of three successive terms from the digital time series are applied to a non-linear processor 4 which, in a manner described in detail below, calculates from the three terms instantaneous values of each of the three parameters, one or more of which may convey the modulation. The calculated values may then be output in numerical form for further processing or may be used directly to reconstitute the original signal.

In order to optimise the output signal to noise ratio the values of the modulated parameter or parameters are low pass filtered with a filter bandwidth equal to that of the original modulating signal (if this is known). If a reduction in data rate is required the filter system 5 used for this purpose may include a decimator. The A/D converter 3, the filter and decimator 5 and non-linear processor 4 may conveniently be formed as part of a single suitably programmed microprocessor which for applications requiring a high processing speed has a parallel processing bit slice architecture.

To determine the instantaneous value of the modulated parameter, e.g. frequency in the case of an FM signal, a process of sinusoidal curve fitting is used. It is assumed that any 3 consecutive points of the digital time series, representing the FM signal, can be associated with a function of the form:-

$$S_i = A \cdot \cos(W \cdot T \cdot i + \Phi) \quad (1)$$

where A is the amplitude, assumed constant over any 3 samples

W is the instantaneous radian frequency

T is the sample interval

i is an integer

Phi is an arbitrary phase offset

For any 3 samples, of unit sample rate, one can state,

$$S_0 = A \cdot \cos(\Phi)$$

$$S_1 = A \cdot \cos(W + \Phi)$$

$$S_2 = A \cdot \cos(2W + \Phi)$$

and

$$(S_0 + S_2) / (2 \cdot S_1) = \cos(W)$$

thus

$$f = (1/(2\pi)) \cdot \arccos((S_0 + S_2)/(2 \cdot S_1)) \quad 0 < f < 0.5$$

or

$$f = -(2/\pi) \cdot \arcsin((S_0 + S_2)/(2 \cdot S_1)) \quad -1 < f < 1 \quad (2)$$

where f is the instantaneous frequency,

acs and asin are the arc cosine and arc sine functions respectively. The next value for f is calculated from terms  $S_1$ ,  $S_2$ ,  $S_3$  and a general term  $f_k$  is given by

$$f_k = -(2/\pi) \cdot \arcsin((S_k + S_{k+2})/(2 \cdot S_{k+1}))$$

where  $S_k$ ,  $S_{k+1}$ ,  $S_{k+2}$  are successive general terms of the digital time series. Values for  $f_k$  are calculated at

a rate equal to the sampling rate of the digital time series. Similarly expressions for general values of the amplitude,  $A_{K1}$  and phase,  $\phi_{iK}$ , in terms of sample values  $S_K$  may be derived. Making the following substitutions,  $X_K = S_{K+1}^2 - S_K S_{K+2}$

$$Y_K = \frac{S_K + S_{K+2}}{2S_{K+1}}$$

the expressions for the three parameters become,

$$f_K = (-2/\pi) \arcsin(Y_K)$$

$$A_K = \left( \frac{X_K}{1 - Y_K^2} \right)^{\frac{1}{2}} \quad (3)$$

$$|\phi_{iK}| = \left| \arcsin \left( \frac{S_K}{A_K} \right) \right| \quad (4)$$

with the sign of  $\phi_i$  given by the sign of  $(S_K Y_K - S_{K+1})$

Since the expression for frequency involves a division by the sample  $S_i$ , it is necessary to include default expressions to accommodate the possibility of  $S_i = 0$ . Two default expressions for instantaneous frequency have been considered for the case  $S_i = 0$ . These are:-

$$f_i = f_{i-1}$$

and

$$f_i = f_{i-1} + (f_{i-1} - f_{i-2})$$

where

$f_i$  refers to the frequency at time  $i$ .

In addition, the above expressions may be applied whenever the absolute value of  $S_i$  is below some threshold value. In the absence of companding digital samples which are small in magnitude compared to the peak value will be least accurate due to quantisation error. It may therefore be preferable to use one of the default expressions for frequency instead of eqn. (2), when the sample magnitude is below a fixed threshold.

In order to evaluate the performance of a method in accordance with the present invention a microcomputer was programmed to generate a frequency modulated signal together with appropriate forms of noise and then to demodulate the signal using the method of the present invention. A flow diagram of the relevant processes is shown in Figure 3. Current capabilities allow both sinusoidal and bandlimited noise functions for the baseband signal form. In the case of noise, frequency emphasis was applied (6 dB/octave) prior to bandlimiting to a maximum modulation frequency equal to 0.02 of the sampling frequency.

An integration process was then applied, followed by multiplication by the 'deviation frequency'. This function was then used to modulate the argument of a sinusoid with a carrier frequency component, generally equal to 0.25 of the sampling frequency.

Having been synthesised the FM signal was demodulated using the 3 point discriminator algorithm, (eqn. (2)), after which the frequency estimates were filtered to maximise the signal to noise ratio (SNR). To obtain a noise figure for the discriminator algorithm itself, the original baseband signal was also filtered by an identical low pass filter. This was then used to compare with the demodulated signal, and by taking the difference between them the 'process noise' generated by the discrimination algorithm was extracted. The SNR was generally estimated over 256 sample points which corresponds to approximately 50 cycles of the highest modulation frequency.

The performance of the algorithm was investigated as a function of SNR at the RF-IF stage and also in relation to parameters intrinsic to the algorithm itself, such as threshold and linear extrapolation.

Figure 4 shows the measured output SNR versus the input SNR for a simple fixed carrier frequency (i.e. zero deviation). Since the model assumes the samples are derived from a sinusoidal function, the output SNR can be in excess of 60 dB when the input SNR is sufficiently high. Figure 5 shows the performance of the discriminator for finite frequency deviation ( $= 0.08$  of sample frequency). The noise power here refers to the noise as measured over the full Nyquist bandwidth and the signal power for the FM

signal is given by  $A^2/2$ , where  $A$  is the amplitude. It can be seen that the output SNR reaches a maximum value of approximately 30 dB when the input SNR is better than 40 dB. The output SNR then falls rapidly below 10 dB input SNR, showing a 'threshold' effect at approximately 15 dB input SNR. Between 10 and 20 dB input SNR the system appears to yield some gain which will be dependent upon the modulation index of the FM signal. The absolute values of the output SNR are characteristic of the particular ratio of  $f_p/f_m$  employed. Where a higher SNR is required this may be obtained by using a higher ratio of  $f_p/f_m$  and a correspondingly higher degree of post demodulation filtering.

The 3 point discriminator algorithm is essentially a continuous sine curve fitting exercise which involves taking 3 successive samples of the FM signal and solving for frequency at each stage. A minimum of 3 points is needed in view of the fact that amplitude, frequency and phase are all unknown values in the expression for a sinusoid. It is therefore apparent that only when the signal is an accurately represented, time invariant sinusoid that the algorithm will yield wholly accurate frequency estimates. In the evaluation the effects of finite quantisation error were included by modelling the digitised signal with 8 bit resolution. With this constraint, it will be appreciated that when the amplitude of the 3 samples is small, relative to the maximum amplitude, (eg. at or near zero crossings), the frequency estimates, given by the algorithm, will be least accurate, tending to infinite values which can only be limited to permissible maximum deviation.

In an attempt to optimise the algorithm in the presence of quantisation error, a threshold test is placed upon the central sample. Provided the value of the sample is greater than the threshold the expression for frequency is evaluated. If it is less, one of a number of options is possible. The simplest option is to select the previous estimate of frequency as the new estimate. Figure 6 shows the output SNR as a function of the threshold. A value for the threshold of 64, (in a maximum of 127) was found to be optimal and with sign/magnitude representation of the sample offered the further advantage that only the most significant bit of the magnitude word needed to be tested to determine the appropriate procedure for evaluating the frequency estimate.

Two versions of the algorithm were examined to test the performance when the central sample was less than the threshold. These were:-

- 1) Set frequency equal to previous estimate
- 2) Set frequency equal to previous estimate + previous gradient

Figure 7 shows the effect on the output SNR of the two algorithms. Normally one would expect algorithm 2) to be superior. However, since the frequency estimates are subject to process noise, linear extrapolation can give rise to very large errors and simulation has shown that setting the new estimate equal to the previous estimate is a better option. In addition, linear extrapolation would require further tests to be applied to the demodulated signal to prevent the frequency estimates from exceeding reasonable limits.

In order to estimate the dynamic range of the output of the discriminator, sinusoidal functions were applied to the input with 'infinite' SNR. Under such circumstances, the only source of noise was the quantisation noise for 8 bit representation. The noise component of the discriminator output was then extracted by subtracting the actual frequency from the estimated frequencies (over a period corresponding to 256 sample points). A value for the rms error in amplitude could then be deduced, and since the maximum amplitude at the discriminator output was unity, the dynamic range could be expressed in dB's as  $-20 \cdot \text{LOG}(\text{rms error})$ . Rms errors were deduced for a range of phase offsets in the sampled time series and the minimum value selected to represent the maximum dynamic range for the discriminator process. Figure 8 shows the results for a range of frequencies offset relative to half Nyquist. To accommodate a dynamic range of approximately 60 dB the output of the discriminator should have a 'word size' of 12 bits.

With regard to the demodulation of frequency multiplexed signals it is usually desirable to ensure equal SNR in all channels of the multiplexed group. To determine the behaviour of a method in accordance with the present invention in respect of this requirement the "process noise" was monitored as a function of frequency. Since the modulating Gaussian noise signal is linearly frequency weighted the process noise should exhibit the same linear frequency dependence in order to maintain the same SNR in all the frequency multiplexed channels which constituted the modulating signal. Figure 9 indicates a levelling-off of the process noise towards low frequencies: as a result channels placed at the lower end of a multiplexed group would be expected to suffer from a greater proportion of process noise. However it is found that the form of the process noise spectrum is dependent upon the choice of threshold level so that in applications where this trend in the process noise would otherwise present significant problems the performance of the modulator may be optimised by choosing a suitable value for the threshold level to balance the requirements of the total SNR level and the uniformity of distribution of the noise.

**Claims**

1. A method of demodulating a signal comprising a high frequency carrier signal having at least one parameter modulated by a low frequency modulating signal, comprising digitising the signal and applying  
5 terms from the resulting digital time series to a non-linear processor to produce an output representative of the instantaneous value of the modulated parameter characterised in that the signal is translated to a low intermediate frequency and subsequently sampled at a correspondingly low sampling rate.

2. A method according to claim 1, in which the non-linear processor produces an output representative of the instantaneous value of the modulated parameter by a process of sinusoidal curve-fitting.

10 3. A method according to claim 1 or 2, in which at least three successive terms of the digital time series are applied to the non-linear processor to produce the output representative of the instantaneous value of the modulated parameter.

4. A method according to claim 3 further comprising taking one further term from the digital time series and applying the one further term and all but the first of the at least three successive terms to the non-linear  
15 processor to produce a further output representative of the next instantaneous value of the modulated parameter, and so on.

5. A method according to claim 1, in which the non-linear processor produces an output representative of the instantaneous value of the modulated parameter by performing a Fast Fourier Transform.

20

25

30

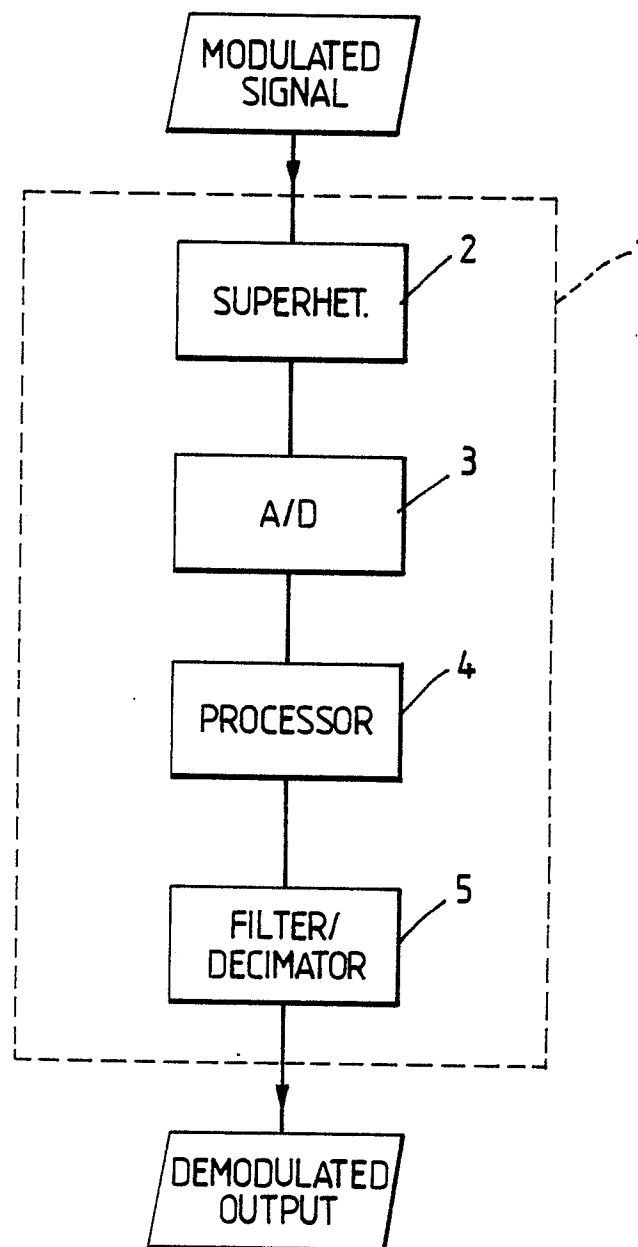
35

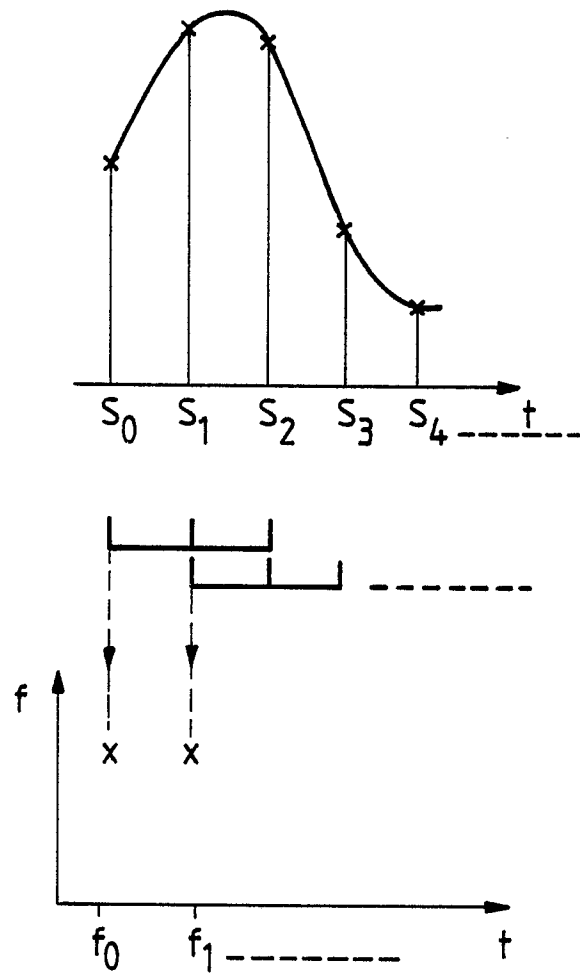
40

45

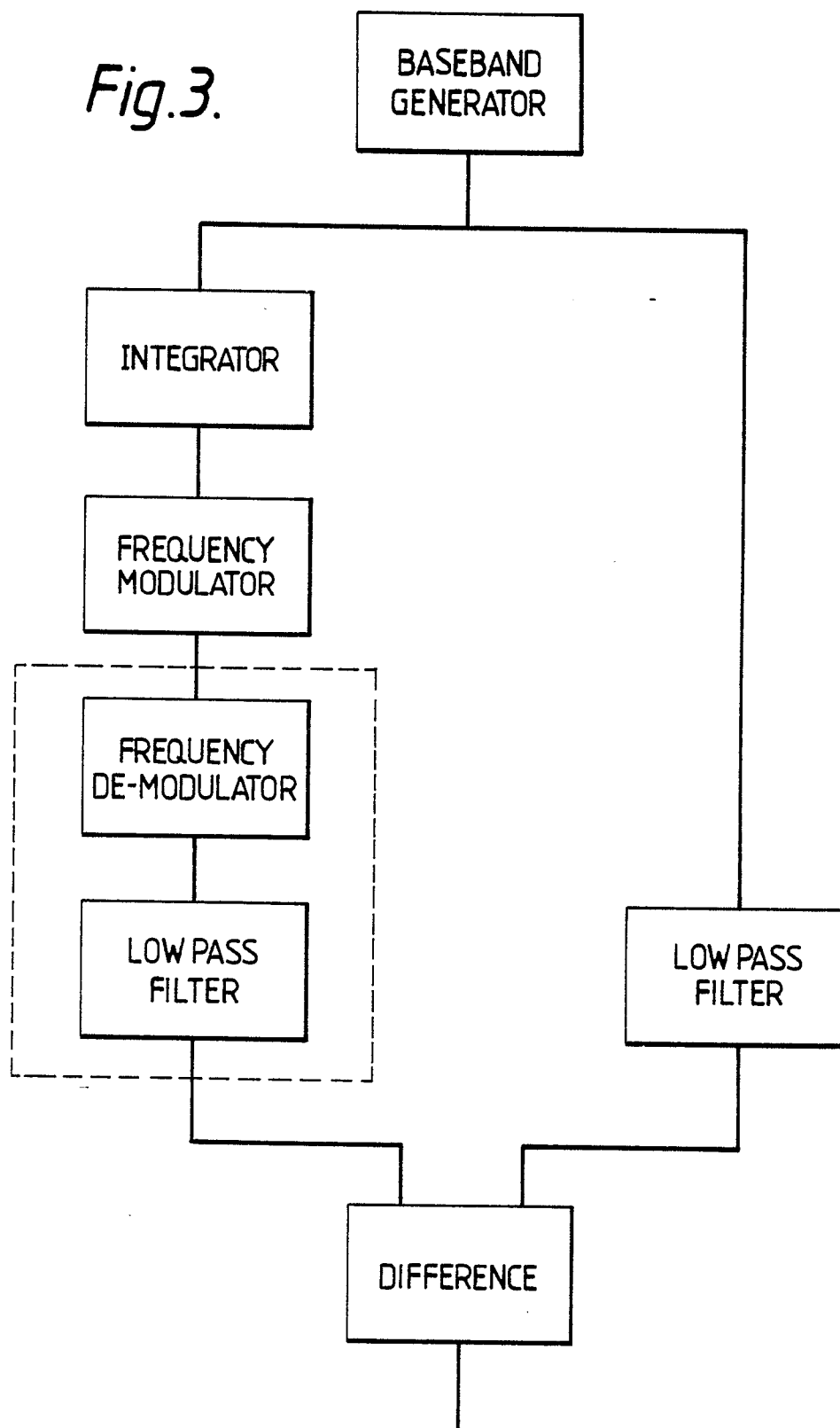
50

55

*Fig.1.*

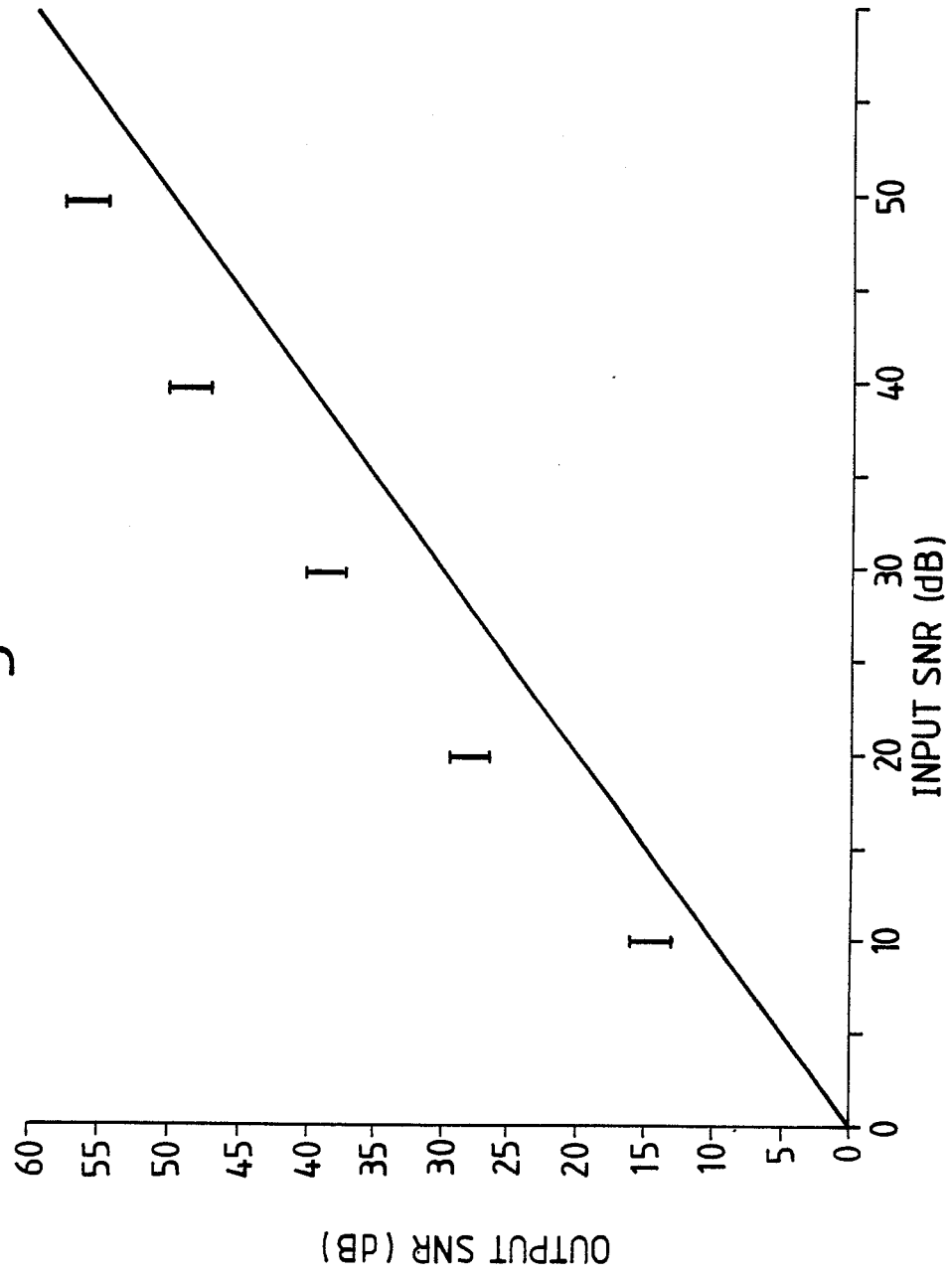
*Fig.2.*



*Fig.3.*

FLOW DIAGRAM OF SIMULATION PROGRAM

Fig.4.



OUTPUT SNR VERSUS INPUT SNR FOR ZERO DEVIATION

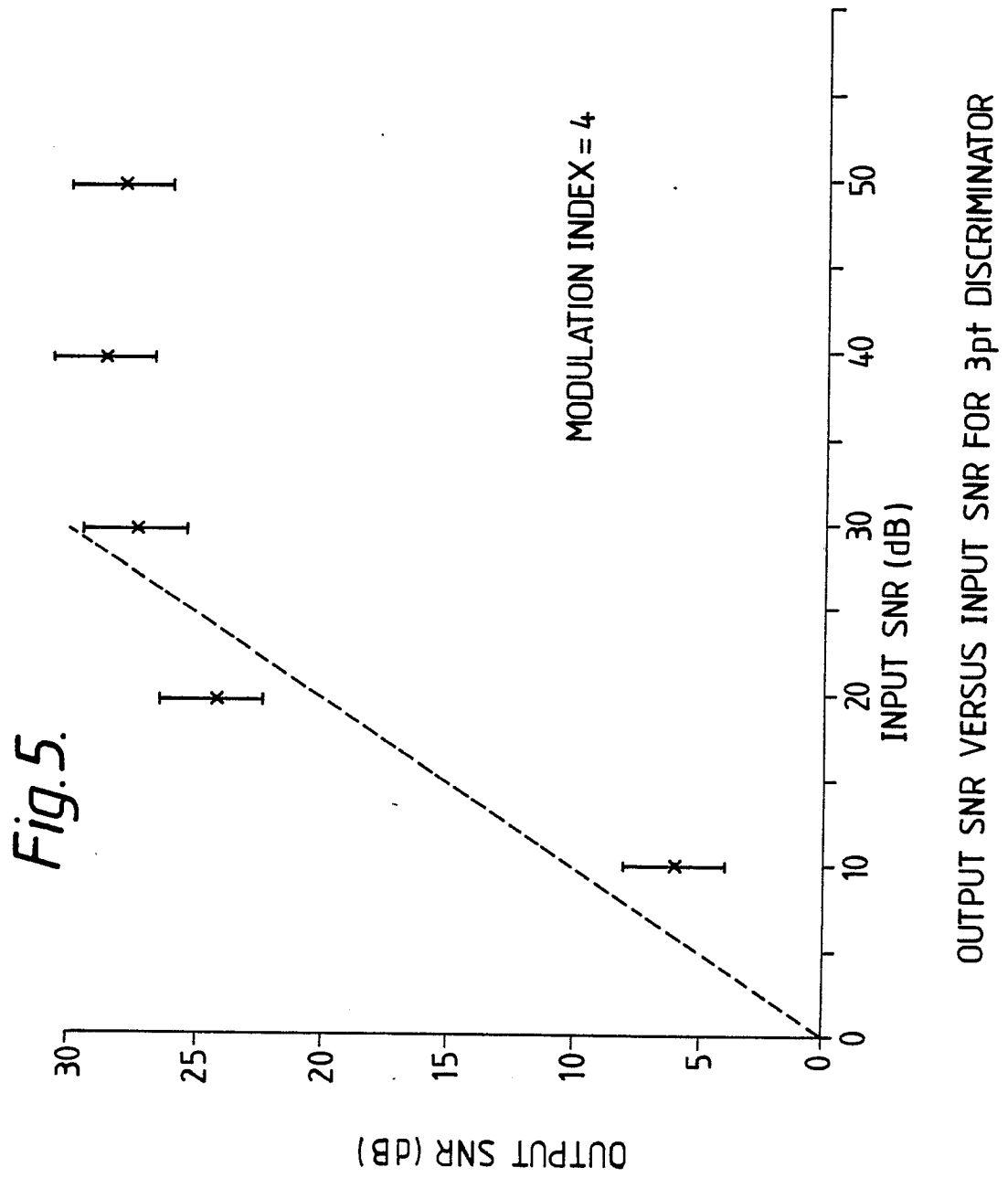
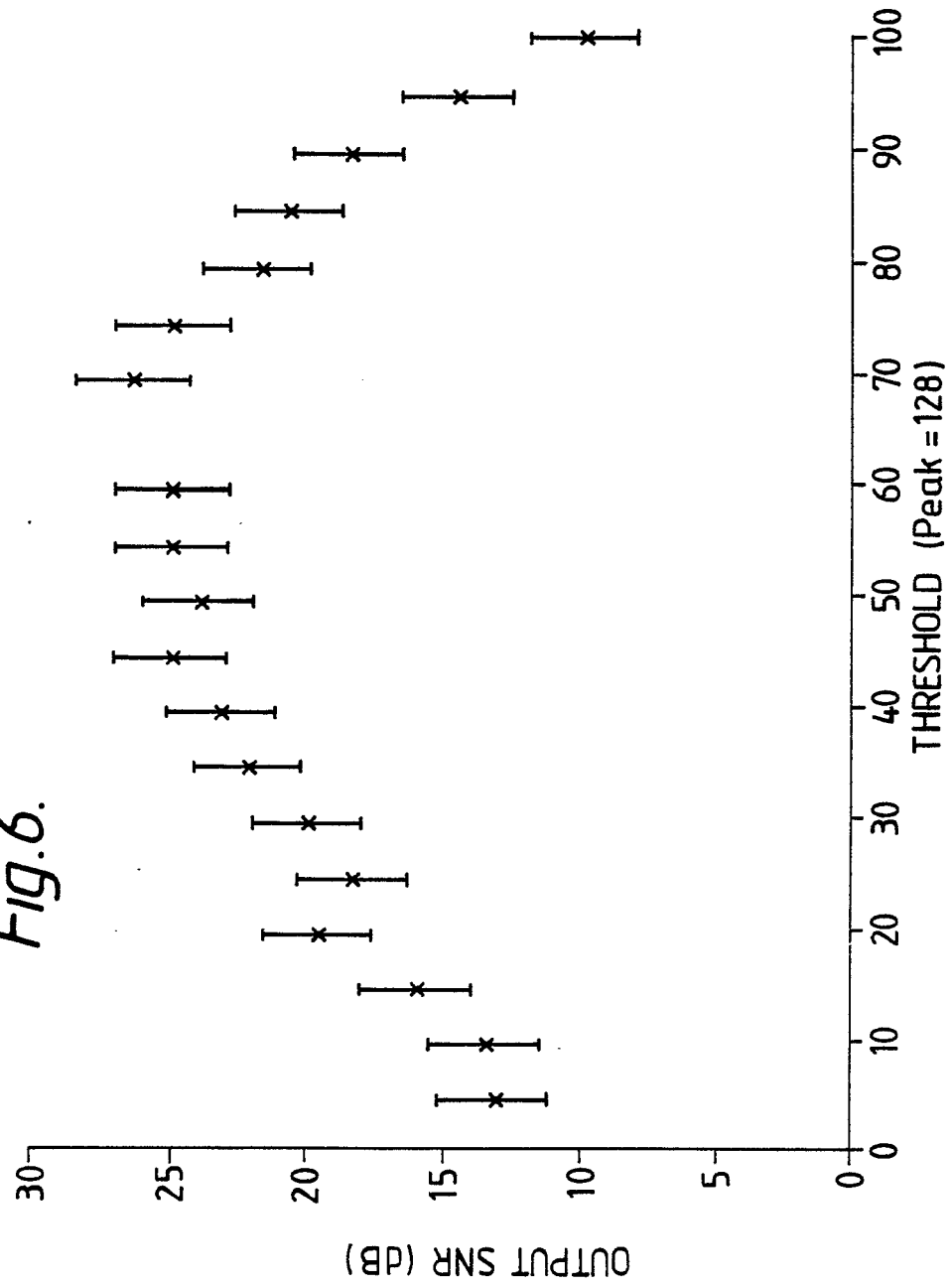
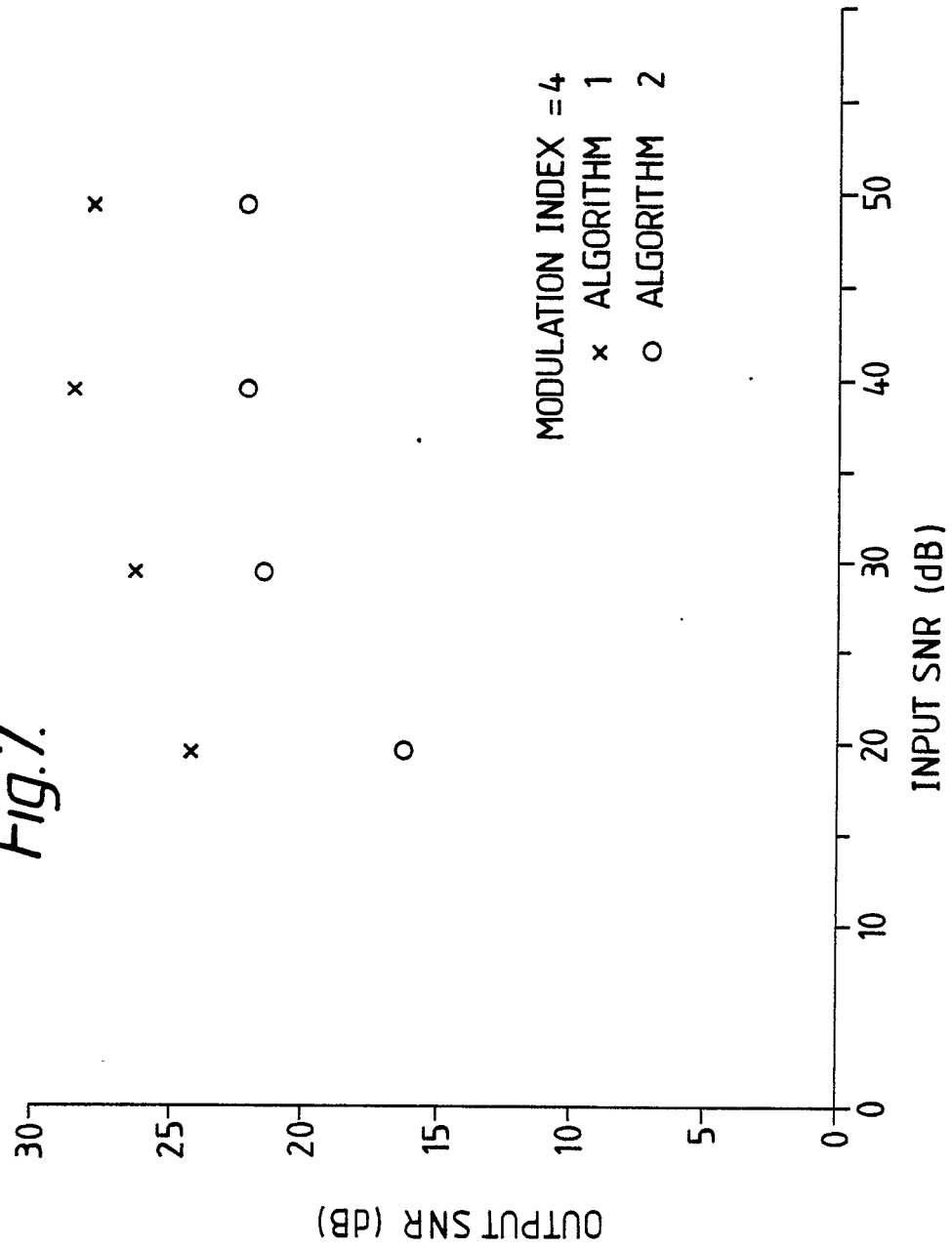


Fig. 6.



OUTPUT SNR VERSUS THRESHOLD FOR 3pt DISCRIMINATOR

Fig. 7.



OUTPUT SNR VERSUS INPUT SNR FOR 3pt DISCRIMINATOR

